

Analyzing interactions between cue-guided and place-based navigation with a computational model of action selection: Influence of sensory cues and training

Laurent Dollé¹, Denis Sheynikhovich^{1,2}, Benoît Girard¹, Balázs Ujfalussy³,
Ricardo Chavarriaga⁴, and Agnès Guillot¹

¹ Institut des Systèmes Intelligents et de Robotique UPMC-Paris 6, CNRS UMR 7222, 4 place Jussieu, 75252 Paris cedex 05-France

² Laboratoire de Neurobiologie des Processus Adaptatifs UPMC-Paris 6, CNRS UMR 7222, 9 quai St. Bernard, 75005 Paris-France

³ KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, Konkoly-Thege Miklós út 29 – 33. Budapest, H-1121, Hungary

⁴ CNBI, Center for Neuroprosthetics, EPFL, Station 11 CH-1015 Lausanne, Switzerland

Corresponding author: laurent.dolle@isir.upmc.fr

Abstract. The hypothesis of multiple memory systems involved in different learning of navigation strategies has gained strong arguments through biological experiments. However, it remains difficult for experimentalists to understand how these systems interact. We propose a new computational model of selection between parallel systems involving cue-guided and place-based navigation strategies allows analyses of selection switches between both control systems, while providing information that is not directly accessible in experiments with animals. Contrary to existing models of navigation, its module of selection is adaptive and uses a criterion which allows the comparison of strategies having different learning processes. Moreover, the spatial representation used by the place-based strategy is based on a recent hippocampus model. We illustrate the ability of this navigation model to analyze animal behavior in experiments in which the availability of sensory cues, together with the amount of training, influence the competitive or cooperative nature of their interactions.

1 Introduction

Animal experiments demonstrate that parallel memory systems, assumed to support the learning of cue-guided and place-based navigation strategies, favour separate sets of sensory cues [1]. According to several studies, the first system, mediated by dorsolateral striatum (DLS), mostly uses proximal cues and the second system, mediated by hippocampus (Hc) and prefrontal cortex (PFC), encodes configurations of distal cues [2]. The spatial representation built from these distal cues is often termed “cognitive map” [3]. Both systems are supposed to interact

competitively or cooperatively, depending on the circumstances. Competition between the systems happens if the inactivation of one enhances the learning of the other, and cooperation if the learning of one compensates the lack of the other (e.g., [4]). Such interactions are influenced by both external and internal factors [5]. Although behavioral and neurophysiological studies provide valuable information about interactions between strategies and their potential biological substrates, the mechanisms underlying these interactions are not clear [6]. This is mainly due to the difficulty of knowing exactly which strategy is chosen at specific moments of the experiments.

We present here a computational model of navigation that provides direct information on rats' behavior in a Morris maze paradigm in which interactions between cue-guided and place-based strategies were shown to be influenced by the type of landmarks and the amount of training [7]. It supplies possible explanations of strategy selection mechanisms that can produce competition or cooperation and makes it possible to estimate the influence of sensory cues on strategy selection. The model is based on the assumptions (i) that both strategies are mediated by separate navigation "experts" that learn in parallel (as proposed by previous computational models [8,9,10]); (ii) that the selection mechanism continuously updates its estimation of the efficiency of both strategies (as in [8,9,10]); (iii) that the learning of both systems are of different nature: cue-guided strategies rely on a procedural "stimulus-response" learning implemented as a TD algorithm while place-based strategies rely on a graph-search algorithm – not dependent on the reinforcement learning framework [10] – that is more flexible and faster to relocate a goal. Point (iii) constitutes the novelty of the model, which led to a major issue: finding the relevant "common currency" allowing to compare the efficiency of strategies having different learning processes. Another novelty is the integration of a recent hippocampus model [11] that computes the spatial representation used by the place-based strategy.

Section 2 describes the model of strategy selection; Section 3 the experimental protocol and the simulation procedure; Section 4 reports the results of computer simulations reproducing the animal data; Section 5 discusses the results in relation to other experimental and computational works and outlooks future work.

2 The Model of Strategy Selection

In our model "Taxon" and "Planning" experts represent DLS and Hc-PFC dependent memory systems, respectively. During navigation, they propose a direction for the next movement according to either visual input (Sensory Cells for the Taxon expert) or the estimated location (Planning graph, built from Place Cells, for the Planning expert). Sensory Cells are learned from the vision of the intra-maze landmark, Place Cells from the vision of the extra-maze landmarks. In addition, a third expert, Exploration, proposes a direction of movement randomly chosen between 0 and 2π (Fig. 1a). The movement actually performed by the animat is determined by a gating network which selects one of the experts

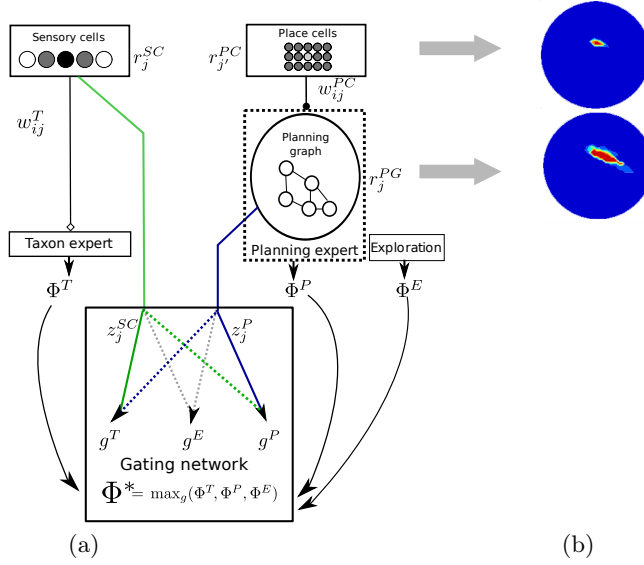


Fig. 1. (a) Model overview (see text for details). SC: Sensory Cells, PC: Place Cells, PG: Planning Graph, T: Taxon expert, P: Planning expert, E: Exploration Expert. Φ^* is the final direction of movement. The Gating network learns to choose the right expert, according to their performance. (b) Example place fields of Place Cells (top) and Graph nodes (bottom), the brighter the more the cell or node are active.

to take control over behavior on the basis of previous performance (Fig. 1a). At each timestep, all the experts learn from the previously executed action, irrespective of which expert was responsible for it (Taxon, Planning or Exploration).

Taxon expert. For the purpose of the protocol reproduced below, we design here the Taxon expert as a “guidance” strategy – approaching a hidden target marked by a landmark located on a certain distance from it [7] – but other forms of Taxon (e.g., beacon approach) can be encoded. The visual input is encoded in a population of $N_{SC}=100$ Sensory Cells (SC) in which the activity of cell i signals presence or absence of the landmark in the direction ϕ^i from the animat. The motor response to the visual stimulus is encoded by $N_{AC} = 36$ Action Cells (AC), so that one action codes for a direction every $2\pi/36$. The learning is done by a Q-learning algorithm adaptation [12], so that the action space is continuous: in the update rule, instead of reinforcing the only action that is chosen, a Gaussian activity profile tuned around the selected action allows the closer actions to update their weights in the same direction. This activity profile is possible since all actions are direction movements, and therefore comparable. The activity in the whole AC population is interpreted as a population code for

the direction ϕ^T of the next movement of the animat, proposed by the Taxon expert. Details of the computations are given in [9].

Planning expert. The Planning expert, inspired by the model of [13], first learns a topological representation of the environment in a reward-independent manner (*map building phase*) and then uses this representation to remember the goal location and to plan an optimal path towards it (*goal planning phase*). The *map building phase* is run during pretraining sessions. The map takes as input the activity of a population of Place Cells. This population is provided by the hippocampus model of [11] consisting of two different neural networks, simulating the enthorinal cortex (EC) and the dentate gyrus (DG). EC modelling is based on the recently discovered grid cells, which are receptive to specific spatial frequency and orientation, and therefore are good candidates for implementing path integration and other navigation strategies like, e.g., praxis. The grid cells are appended to a vector of 100 gray units representing the sight of the extra-maze landmarks. EC cells are then fed to the DG, which produces Place Cells by means of a Hebbian learning. Then a sparse representation is computed with a filter function that only keeps a few cells active (Fig. 1b, top) and sets the others to zero. Detailed computations can be found in [11]. The final activation is then normalized and processed by the Planning expert to build the nodes of the planning graph (PG) (Fig. 1b, bottom). For that, a pool of 100 nodes is connected to the DG at random synaptic weights, and the sparse learning used in the DG is replicated during the map building phase.

First the firing rate of a PG node j is computed as follows:

$$r_j^P = f_j\left(\sum_i W_{ij}^{(DG,PG)} DG_i, s^{PG}\right), \quad (1)$$

where W_{ij} is the synaptic weight linking the DG place cell i to the PG node j and $f_j(x, s^{PG})$ is the same non-linear function as in the DG, returning a sparse encoding of x , with a sparseness level of s^{PG} .

The synaptic weights are learned following a Hebbian rule, similar to the one used for learning the DG output:

$$\Delta W_{i,j}^{(DG,PG)} = \alpha^{(DG,PG)} r_j^P (DG_i - W_{i,j}^{(DG,PG)}) \quad (2)$$

A link between nodes N_i and N_j stores the allocentric direction of movement required to move from one node to the other. The *goal planning phase* begins when the goal position is found, the closest node being set to the delivered reward value. Then, given the PG, the optimal path to the goal is determined by the bio-inspired activation-diffusion mechanism [14] based on Dijkstra's algorithm for finding the shortest path between two nodes in a graph [15]. At each timestep, the Planning expert proposes the corresponding direction. If the goal position is not known, a random direction is proposed.

Strategy selection. Each expert computes at the same time its own proposition of movement. The time spent by the Planning expert to compute the path is not taken into account as a potential cost in our selection mechanism. The

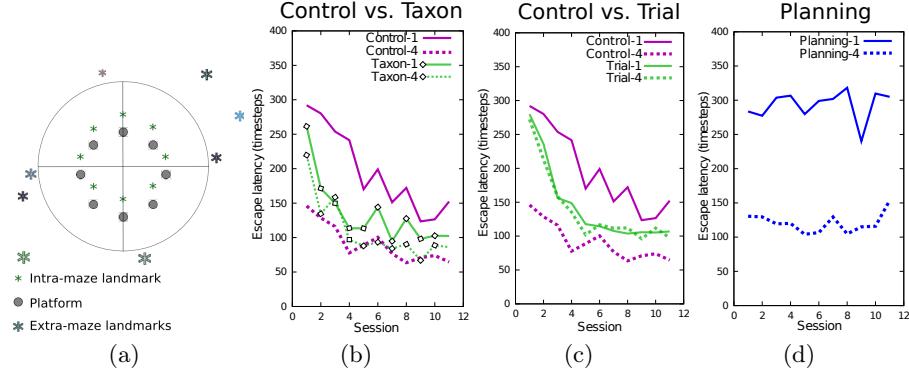


Fig. 2. (a) Simulated environment with one intra- and eight extra-maze landmarks. Simulated results of (b) Control vs Taxon groups. (c) Control vs Trial groups. (d) Planning group only, in the same protocol as the Control group.

gating network selects at each timestep which of the Taxon, Planning or Exploration experts (T, P and E) will control the future movement, on the basis of candidate directions Φ^k of movement. It consists of three units $k \in \{T; P; E\}$, each corresponding to an expert. The activity g^k of expert k is its “gating value”. They are computed as the weighted sum of both SC and PG inputs by synaptic weights z^k :

$$g^k(t) = \sum_{j=1}^{N_{SC}} z_j^k(t) r_j^{SC}(t) + \sum_{j=N_{SC}+1}^{N_{SC}+N_P} z_j^k(t) r_j^P(t), \quad (3)$$

where z_j^k is the connection weight between the unit k of the gating network and input unit j of the experts. A winner-take-all scheme then chooses the next movement direction $\phi^{k'}$:

$$\phi^{k'}(t); k' = \operatorname{argmax}_i (g^i(t)) \quad (4)$$

The gating values connection weights are adjusted using the same Q-learning algorithm as the Taxon expert, except that the update is modulated by the angular difference between the proposed orientation and the one actually chosen, so that the closer an orientation is from the chosen one, the stronger is its update.

We then evaluate different versions of the model – some corresponding to the simulation of lesioned animals – in the experimental paradigm described below.

3 Experimental Paradigm and Simulation Procedure

In the experiment of [7], two groups of rats – intact (Control) and Hc-lesioned – learned to find the location of a hidden platform in a Morris water maze surrounded by several room landmarks. The platform was cued by a visible

landmark located in the pool at a certain distance northward to the platform. In a first experiment, for both Control and Hc groups, the platform and landmark were moved to one of eight predefined locations at the start of eleven sessions, where they stayed for four trials. In a second experiment, intact rats were tested in the same apparatus, but the platform and landmark were moved from trial to trial rather than from session to session (Trial group). The same sensory cues were available for both tasks. Then the only difference was the possibility or the incapacity of training one or the other strategy system within each session.

Considering their results (Fig. 3 and 5 in their paper), the authors hypothesized (i) that, contrary to Hc group which could only learn and use a cue-guided strategy, Control and Trial groups would be able to acquire both strategies, thanks to the presence of cues inside and outside the pool; (ii) that all groups would be able to achieve the tasks (i.e., with one or both strategies available); (iii) that moving the platform from trial to trial would have the same effect on intact rats as a hippocampal lesion: rats of Trial group would indeed exclusively rely on a cue-guided strategy, as they would not be able to sufficiently train to refresh their cognitive map between two successive platform displacements. Our model will test each of these hypotheses by analyzing the interactions between these control systems.

The simulated water maze, rat, reward location and landmark are represented by circles of 200 cm, 15 cm, 10 cm, and 20 cm in diameter, respectively. The reward location is always 20 cm south from the landmark. 8 extra-maze landmarks are placed at different distance of the walls (30 cm to 50 cm), simulated by Gaussian visual stimuli of different widths (Fig. 2a). Three groups are simulated, corresponding to those tested in the actual experiment: the Control and Trial groups of intact rats are simulated with Taxon, Planning and Exploration experts; the Hc-lesioned group, henceforth called Taxon group, is simulated with Taxon and Exploration experts only. The same training protocol as in [7] is applied. We add a fourth group (Planning group, which could correspond to DLS-lesioned animals), not present in the original experiment, which is simulated with Planning and Exploration experts only, tested with the same protocol as the Control group. For each group, 100 sets of experiments were performed.

The performance of the Control, Trial and Taxon groups is statistically assessed by comparison of their mean escape latencies (number of timesteps to reach the goal), within and across sessions, in the 1st and 4th trials of a session, using signed-rank Wilcoxon test for matched-paired samples. Between-group comparison is performed using a Mann-Whitney test for non matched-paired samples. Two measures quantize the animat’s behavior: *goal occupancy rate* (number of times the animat visits a rewarded zone divided by the total trajectory length); *trial selection rate* of an expert (averaged number of times the expert is selected over the total length of the trajectory). The influence of sensory inputs on the selection of the strategies is assessed by comparing synaptic weights from the SC and from the PG nodes to the units of the gating network, which correspond to the relevance the strategies.

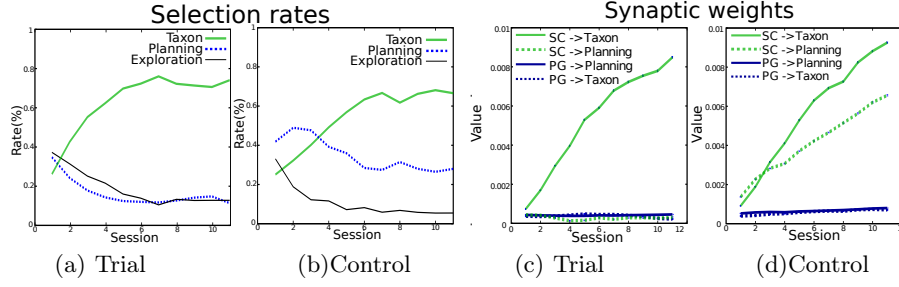


Fig. 3. Trial selection rates of experts averaged over four trials (one session) for (a) Trial and (b) Control groups. Evolutions of synaptic weights in the gating network across sessions for (c) Trial and (d) Control groups show the relative influence of all sensory cues on the selection.

4 Results

Test of performance of separate experts Separately trained on a simple task in which the animat had to escape to a visible and fixed platform, both experts differ in their learning processes. The Taxon expert learns slower than the Planning expert (mean of escape latencies from the 1st to 4th trials for Taxon = 235.04; for Planning = 160.82) but has a better performance when the training becomes intensive (mean of escape latencies from the 40th to 44th trials for Taxon = 21.08; for Planning = 83.18). These differences are due to the sparsity of the Planning Graph nodes, which allows this expert to quickly locate the new goal location (as encoded by existing nodes), but no further learning is done.

All groups are able to learn the tasks. As Fig. 2b and c attest, the model reproduces the main characteristics of the original experimental results. Indeed all groups achieve the tasks, as all escape latencies decrease across sessions ($p < 0.001$). In accordance with [7], the cue-guided strategy is mainly responsible of this improvement: in Control and Trial groups, the Taxon selection rate dominates the others and increases throughout the sessions (Fig. 3a,b). In all gating networks, the evolution of weights between sensory cues and experts also reflects the growing influence of SC on Taxon expert (Fig. 3c,d).

No interaction between experts in Trial group. The similar performance of Taxon and Trial groups ($p = 0.08$; learning across, but not within sessions) suggests that the Taxon expert mainly controls the behavior of Trial animats (Fig. 2b, c). This is confirmed by Fig. 3a, showing the rapid and huge increase of the selection rate of this expert during the experiment, while the others remain very low. This is also illustrated by typical trajectories, showing that the Taxon expert mainly leads the animat near the platform (Fig. 4a, b). The results of the Taxon group indicate that its learning process is not flexible enough to quickly improve within session. Thus this explains why Trial group behaves

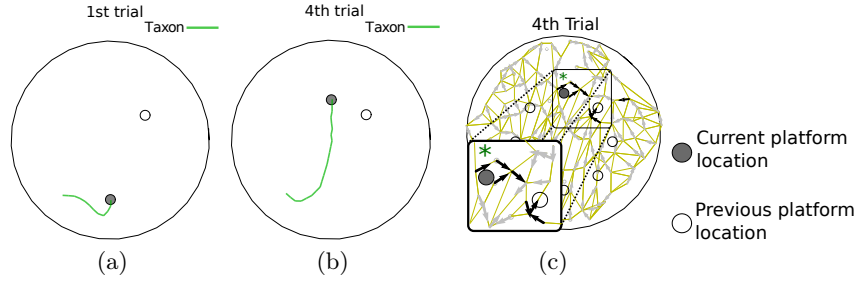


Fig. 4. (a-b) Typical trajectories of Trial animats, corresponding to beginning (a) and end (b) of session; (c) Navigational map of Planning expert in the Trial group (i.e., learned directions of movement for each spatial node) at the end of the same session. The zoom depicts that, during a new trial, the planning experts keeps the memory of the previous platform location.

similarly: the possibility of training or not during four trials does not change the performance.

The occupancy rates near the current platform (fig 5a, 1st and 4th trials) are significantly lower in the Trial than in the Taxon groups. This may suggest a negative influence of the Planning expert on Trial animats. Yet the evolution of weights in the gating network, with increasing links from SC to the Taxon expert only (Fig. 3c), minimizes the importance of this result: it clearly indicates that the Planning expert does not play any role in this task. Nevertheless, this expert continuously builds a navigational map in which it stores the successive platform locations (Fig. 4c), then the animat could use it if required by a new situation.

Training within session produces both competition and cooperation. As Fig. 2b shows, the possibility to train within session modifies significantly the behavior of Control animats, both at the beginning and at the end of sessions. Similarly to Control animals in the original experiment, they perform worse at the beginning and better at the end of one session than Trial and Taxon groups. The role of Taxon in Trial group proves that these differences should come from the Planning expert. Fig. 3b confirms that Control animats select it more often than Trial ones (especially at the beginning of the experiment). The fast learning of this expert – also favored by the parallel learning of the efficient Taxon expert (see the performance of Planning alone, Fig. 2d) – indeed allows the Control group to quickly memorize the position of the platform in the cognitive map. During each session, both experts cooperate and enhance the overall performance: the escape latencies are lower (Fig. 2b) and the occupancy rates near the current location higher (Fig. 5a, 4th trials) than Trial group. The gating network reflects this synergistic interaction by increasing the weighting of SC to both Taxon and Planning experts (Fig. 3d). The nature of their cooperation could be deduced from the navigational maps of both experts at the

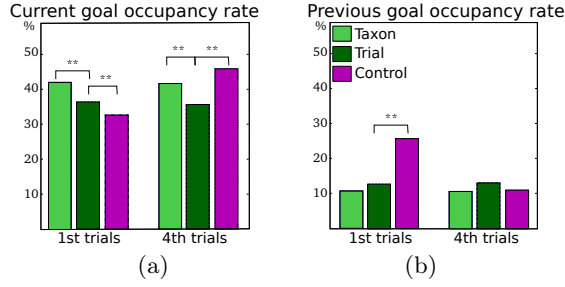


Fig. 5. Occupancy rates near the (a) current and (b) previous platform location.

end of a session showing that the Taxon expert drives the animat southward the landmark (Fig. 6a) whereas the Planning expert leads it to the platform location (Fig. 6b). A typical trajectory illustrates this hypothesis (Fig. 6c).

At the beginning of a new session, as the gating network grants more confidence to the Planning, this expert has a stronger tendency to be selected. As a consequence, both Planning and Taxon experts compete, driving the animat respectively towards the previously memorized (thus wrong) platform location or the current one (Fig. 6d). Fig. 5b shows the significant differences of occupancy rates near the previous platform location between Control, Trial and Taxon groups during the 1st trials, explaining the worse performance of the Control group. The selection rates of Planning and Taxon experts in Fig. 3b however show that this competition progressively decreases across sessions. Indeed the Taxon expert takes more and more control over Planning (negative correlation between Taxon and Planning selection rates: $r = -0.91$). Exploration is used until both experts have switched (i.e., until Taxon becomes sufficiently relevant, Fig 3b).

5 Discussion and Conclusion

The model of selection between Taxon and Planning navigation strategies presented here is based on the theory of parallel control systems in the rat brain. The place-based strategy uses a graph-search algorithm using the propagation of the reward signal to find the shortest path to the goal. The graph is learned on-line in pretraining sessions using the activities of learned place cells, and positions the location of the hidden goal. The cue-guided strategy uses a TD learning rule to approach the hidden goal marked by a landmark. The strategy selection is performed by a gating network that learns to predict, also using a TD-learning rule, the most successful strategy given current sensory input. Previous computational models of navigation rely on similar learning modes (e.g., dependent on the RL framework) for different navigation strategies [8,9,10], and the strategy selection mechanism is usually non-adaptive [16,17]. In contrast, our model al-

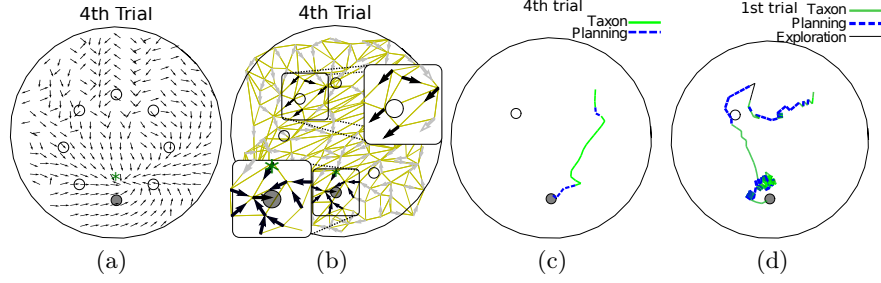


Fig. 6. Control group. Example of navigational maps of Taxon (a) and Planning (b) experts (learned directions of movement for each sample location or each spatial node) at the end of sessions. All possible platform locations are positioned, the current one is in grey. Taxon map is pointing the current landmark; Zooms in Planning map depict that the map is pointing the current platform position, but not the previous one. (c) Typical trajectory at the end of sessions (4th trial). (d) Typical trajectory at the beginning of sessions (1st trial) showing the planning expert leading the animat towards the wrong location.

lows for adaptive selection among different learning modes (including topological representations), exploiting a general and simple “common currency”.

The model was evaluated in two simulated water maze tasks, in which the same sensory cues (intra-maze and extra-maze landmarks) were available. In one task, the cued hidden platform moves at each trial; in the other, the goal stays at the same place during a four trials session before moving to another location (Trial and Session tasks in [7]). Then what differs is the possibility, or not, to learn each position of the platform. Due to a separation between cooperative interactions (during action learning) and competitive interactions (during action selection), the model was able to assess the relative contribution of different strategies to the observed behavior. In accordance with the results of the original experiment, the selection mechanisms of the model – which did not change over the experiment – could explain why both place-based and cue-guided strategies did not interact in the trial condition, and why they cooperated or competed when they (particularly the place-based strategy) could learn the same platform position during several trials. A substantial contribution of the model concerned the analysis of the influence of different types of sensory cues on strategy selection. From the evolution of synaptic weights between sensory inputs and gating units in our simulations, it made it possible to assess that intra-maze landmark information was predominant for strategy selection in both tasks and moreover that it contributed to the selection of the Planning when this expert was allowed to be trained (see evolution of SC→Planning, Fig. 3d). Indeed, we remind that the gating network updates its selection on the basis of *both* sensory inputs.

We showed the ability of our model to efficiently select navigation strategies in two experiments in which the effects of environmental cues and training

on navigation system interactions were rather simple. In particular conditions, these interaction may be more complex. For example, under the hypothesis that all spatial cues compete for predicting reward, numerous experiments supporting associative theory emphasize blocking (i.e., when a well learned cue predicting reward prevents learning of a novel cue predicting the same reward) or overshadowing effects (i.e., when one cue predicting the reward detracts the learning of another present cue able to predict the same reward) in navigation [18]. Some of these effects are challenged by experiments supporting the cognitive mapping theory (e.g., [3]). Our model is a potential tool for investigating such contradictions.

Besides, the effect of practice on the selection of navigation strategies is far less investigated than the influence of sensory cues. This was particularly done in the experiments of [19], which stressed the complexity of this factor in studying the influence of intensity of training experience during one session on the later use of a specific strategy. In contrast to previous conclusions (e.g., [20]), short or long training, but also various periods of pre-exposure to the task, critically determine when and how a particular strategy emerges from interactions between both memory systems. In the future, our model could be applied to analyze such influences.

However we have to notice that the validation of the model was made in experiments using Morris water mazes, in which every orientation is left opened and not constrained by corridors. A recent paper raises the hypothesis that, in this kind of device, expression of strategy switching may be different from T- or Star-mazes [21]. Then we need to verify if the selection mechanism of our model supports such eventuality. More importantly, the same paper questions the role of DLS, with the assumption that it is not especially involved in ego-centric/response strategy, but in the selection itself, in case of several available strategies. This requires analyzing, in similar protocols as theirs, which bias the specific role attributed to our expert Taxon is susceptible to entail, and in which extent it should be questioned.

Interaction among several spatial memory systems may improve the performance of animals either by speeding learning through cooperation of different strategies, or competitive processes that prevents sub-optimal strategies to be applied. Better understanding of these interactions by computational modelling may also provide a good basis for the design of robots able to cope with a wider range of behavioral tasks.

Acknowledgment This research was granted by the EC Integrated Project ICEA (Integrating Cognition, Emotion and Autonomy, IST 027819).

References

1. White, N., McDonald, R.: Multiple parallel memory systems in the brain of the rat. *Neurobiol Learn Mem* **77** (2002) 125–184

2. Doeller, C.F., King, J.A., Burgess, N.: Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proc Natl Acad Sci U S A* **105**(15) (2008) 5915–5920
3. O’Keefe, J., Nadel, L.: The hippocampus as a cognitive map. Oxford University Press. (1978)
4. Hartley, T., Burgess, N.: Complementary memory systems: competition, cooperation and compensation. *Trends Neurosci* **28**(4) (2005) 169–170
5. Carrillo-Mora, P., Magda, G., Abel, S.: Spatial memory: Theoretical basis and comparative review on experimental methods in rodents. *Behav Brain Res* **203**(2) (2009) 151–164
6. Rich, E., Shapiro, M.: Rat Prefrontal Cortical Neurons Selectively Code Strategy Switches. *J. Neurosci.* **29**(22) (2009) 7208–7219
7. Pearce, J., Roberts, A., Good, M.: Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. *Nature* **396**(6706) (1998) 75–77
8. Dolle, L., Khamassi, M., Girard, B., Guillot, A., Chavarriaga, R.: Analyzing interactions between navigation strategies using a computational model of action selection. *LNAI* **5248** (2008) 71–86
9. Chavarriaga, R., Strosslin, T., Sheynikhovich, D., Gerstner, W.: A computational model of parallel navigation systems in rodents. *Neuroinformatics* **3**(3) (2005) 223–242
10. Uchibe, E., Doya, K.: Reinforcement learning with multiple heterogeneous modules: A framework for developmental robot learning. In: The 4th International Conference on Development and Learning. (2005) 87–92
11. Ujfalussy, B., Eros, P., Somogyvari, Z., Kiss, T.: Episodes in space: A modelling study of hippocampal place representation. *LNAI* **5040** (2008) 123–136
12. Stroesslin, T., Sheynikhovich, D., Chavarriaga, R., Gerstner, W.: Robust self-localisation and navigation based on hippocampal place cells. *Neural Net.* **18**(9) (2005) 1125–1140
13. Martinet, L.E., Passot, J.B., Fouque, B., Meyer, J.A., Arleo, A.: Map-based spatial navigation: A cortical column model for action planning. *LNAI* **5248** (2008) 39–55
14. Burnod, Y.: Organizational levels of the cerebral cortex: an integrated model. *Acta Biotheor* **39**(3-4) (1991) 351–361
15. Dijkstra, E.: A note on two problems in connection with graphs. *Numer. Math.* **1**(269-270) (1959) 269–271
16. Girard, B., Filliat, D., Meyer, J., Berthoz, A., Guillot, A.: Integration of navigation and action selection functionalities in a computational model of cortico-basal-thalamo-cortical loops. *Adapt. Behav.* **13**(2) (2005) 115–130
17. Guazzelli, A., Corbacho, F., Bota, M., Arbib, M.: Affordances, motivation, and the world graph theory. *Adapt. Behav.* **6**(3) (1998) 435–471
18. Leising, K., Blaisdell, A.: Associative basis of landmark learning and integration in vertebrates. *Comp. Cogn. Behav. Rev.* **4** (2009) 80–102
19. Martel, G., Blanchard, J., Mons, N., Gastambide, F., Micheau, J., Guillou, J.: Dynamic interplays between memory systems depend on practice: The hippocampus is not always the first to provide solution. *Neuroscience* **150**(4) (2007) 743–753
20. Chang, Q., Gold, P.E.: Switching memory systems during learning: changes in patterns of brain acetylcholine release in the hippocampus and striatum in rats. *J. Neurosci.* **23**(7) (2003) 3001
21. Botreau, F., Gisquet-Verrier, P.: Re-thinking the role of the dorsal striatum in egocentric/response strategy. *Front. Behav. Neurosci.* **4** (2010) 1–12